OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

G243

Statistics 3 (Z3)

QUESTION PAPER

Candidates answer on the printed answer book.

OCR supplied materials:
- Printed answer book G243
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Thursday 16 June 2011
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the printed answer book and the question paper.

- The question paper will be found in the centre of the printed answer book.
- Write your name, centre number and candidate number in the spaces provided on the printed answer book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the printed answer book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the printed answer book and the question paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the question paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The printed answer book consists of **12** pages. The question paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

- Do not send this question paper for marking; it should be retained in the centre or destroyed.

Section A (45 marks)

1 An advertisement claims that 'brain training' increases scores in word-matching games. A psychologist decides to investigate this claim. A random sample of 10 people is selected and all of them play the same word-matching game. After one month of doing brain training exercises, these 10 people repeat the word-matching game.

Player	A	B	C	D	E	F	G	H	I	J
Score before brain training	62	56	36	52	58	94	76	58	62	64
Score after brain training	76	54	53	47	48	87	91	62	71	83

(i) Explain why the psychologist has used a paired design. **[2]**

(ii) Use the Wilcoxon signed rank test, at the 5% level of significance, to examine the claim. **[9]**

(iii) The psychologist originally planned to carry out a paired sample t test. However, she decided that the distributional assumption required for this test might not have been satisfied. State this assumption and briefly explain why in general, if this assumption is satisfied, it is preferable to carry out a t test rather than a Wilcoxon test. **[4]**

2 The manager of a gym is investigating a claim that, on the whole, women spend more time using the gym per visit than men. The manager selects independent random samples of 50 women and 50 men using the gym. He asks each person to record the time, in minutes, spent on this visit. The times for the 50 women are summarised by $\Sigma x = 3308$, $\Sigma x^2 = 222\,020$. For the 50 men, the sample mean is 62.68 and the sample variance is 134.0.

(i) Calculate the sample mean and sample variance for the women. **[3]**

(ii) Explain briefly why, even though the population variances are unknown, it is appropriate to use a test based on the Normal distribution to investigate the claim. **[1]**

(iii) Carry out a test at the 10% significance level to investigate the claim. State your hypotheses and conclusions clearly. **[11]**

3 A random sample of 11 countries is selected. The adult literacy rate and the percentage of children under 16 who attend school in these countries are shown in the table.

Country	A	B	C	D	E	F	G	H	I	J	K
Adult literacy rate	41	56	90	97	83	95	98	29	59	76	68
Percentage attending school	67	70	78	91	84	94	92	47	75	90	83

(i) Draw a scatter diagram to illustrate these data. **[3]**

(ii) Comment on whether, in view of the scatter diagram, it is appropriate to carry out a hypothesis test based on the product moment correlation coefficient. **[2]**

(iii) Calculate the value of Spearman's rank correlation coefficient. **[5]**

(iv) Using your answer to part **(iii)**, test at the 1% level of significance whether it is reasonable to assume that adult literacy rate and percentage of children under 16 who attend school are positively associated. State your hypotheses and conclusions clearly. **[5]**

Section B (27 marks)

4 Researchers at a crop breeding institute are investigating how the yield of a particular variety of pea is affected by the application of either organic or non-organic fertiliser. They apply three different treatments to a number of plots of peas in a field.

Treatment A:	no fertiliser.
Treatment B:	organic fertiliser.
Treatment C:	non-organic fertiliser.

(i) At least one of the three treatments is a control and at least one is experimental. Identify the control and experimental treatments. **[2]**

(ii) Explain why the researchers apply no fertiliser to one area. **[1]**

(iii) Explain why it would not be sensible to divide the field into three strips and apply one treatment to each strip. **[2]**

(iv) The field is in fact divided into 100 equal-sized plots. 10 of these plots are randomly assigned to each treatment. Explain how 30 plots can be selected at random from 100 plots. **[3]**

Two of the plots with Treatment A (no fertiliser) and three with Treatment B (organic fertiliser) were infected with a disease and the whole of the crop in each of these plots had to be destroyed. The yields, in kg per plot, from the remaining plots were as follows.

Treatment A	231	256	221	289	216	238	241	262
Treatment B	278	294	241	250	290	309	291	

For Treatment A the sample mean is 244.25 and the sample variance is 574.2.
For Treatment B the sample mean is 279 and the sample variance is 612.7.

(v) It is thought that the average yield for the crop with organic fertiliser may be greater than that for the crop with no fertiliser. State the assumptions which are required for a t test to examine whether the mean yields of the crops with organic fertiliser and with no fertiliser appear to be the same. Given that these assumptions are valid, carry out the test at the 5% significance level, stating your hypotheses and conclusions clearly. **[14]**

(vi) The researchers wish to test whether there is any difference in average yield between peas grown with organic fertiliser and those grown with non-organic fertiliser. They carry out a Wilcoxon rank sum test. State appropriate hypotheses for the test. Given that the value of the test statistic is 42 and the sample sizes for organic and non-organic are 7 and 8 respectively, complete the test at the 5% level, stating your conclusions clearly. **[5]**

OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

G243

Statistics 3 (Z3)

PRINTED ANSWER BOOK

Candidates answer on this printed answer book.

OCR supplied materials:
- Question paper G243 (inserted)
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Thursday 16 June 2011
Afternoon

Duration: 1 hour 30 minutes

Candidate forename		Candidate surname	

Centre number						Candidate number				

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the printed answer book and the question paper.

- The question paper will be found in the centre of the printed answer book.
- Write your name, centre number and candidate number in the spaces provided on the printed answer book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the printed answer book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the printed answer book and the question paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the question paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The printed answer book consists of **12** pages. The question paper consists of **4** pages. Any blank pages are indicated.

Section A (45 marks)

1 (i)	
1 (ii)	

1 (ii)	**(continued)**
1 (iii)	

2 (i)

2 (ii)

2 (iii)	

3 (i)

3 (ii)

3 (iii)

3 (iv)

Section B (27 marks)

4 (i)	
4 (ii)	
4 (iii)	

4 (iv)

4 (v)

4 (v)	**(continued)**

4 (vi)	

OCR
RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE

Unit **G243:** Statistics 3 (Z3)

Mark Scheme for June 2011

Oxford Cambridge and RSA Examinations

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of pupils of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, OCR Nationals, Functional Skills, Key Skills, Entry Level qualifications, NVQs and vocational qualifications in areas such as IT, business, languages, teaching/training, administration and secretarial skills.

It is also responsible for developing new specifications to meet national requirements and the needs of students and teachers. OCR is a not-for-profit organisation; any surplus made is invested back into the establishment to help towards the development of qualifications and support which keep pace with the changing needs of today's society.

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by Examiners. It does not indicate the details of the discussions which took place at an Examiners' meeting before marking commenced.

All Examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the Report on the Examination.

OCR will not enter into any discussion or correspondence in connection with this mark scheme.

Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

Marking instructions for GCE Mathematics (MEI): Statistics strand

1. You are advised to work through the paper yourself first. Ensure you familiarise yourself with the mark scheme before you tackle the practice scripts.

2. You will be required to mark ten practice scripts. This will help you to understand the mark scheme and will not be used to assess the quality of your marking. Mark the scripts yourself first, using the annotations. Turn on the comments box and make sure you understand the comments. You must also look at the definitive marks to check your marking. If you are unsure why the marks for the practice scripts have been awarded in the way they have, please contact your Team Leader.

3. When you are confident with the mark scheme, mark the ten standardisation scripts. Your Team Leader will give you feedback on these scripts and approve you for marking. (If your marking is not of an acceptable standard your Team Leader will give you advice and you will be required to do further work. You will only be approved for marking if your Team Leader is confident that you will be able to mark candidate scripts to an acceptable standard.)

4. Mark strictly to the mark scheme. If in doubt, consult your Team Leader using the messaging system within *scoris*, by email or by telephone. Your Team Leader will be monitoring your marking and giving you feedback throughout the marking period.

   An element of professional judgement is required in the marking of any written paper. Remember that the mark scheme is designed to assist in marking incorrect solutions. Correct *solutions* leading to correct answers are awarded full marks but work must not be judged on the answer alone, and answers that are given in the question, especially, must be validly obtained; key steps in the working must always be looked at and anything unfamiliar must be investigated thoroughly.

   Correct but unfamiliar or unexpected methods are often signalled by a correct result following an *apparently* incorrect method. Such work must be carefully assessed. When a candidate adopts a method which does not correspond to the mark scheme, award marks according to the spirit of the basic scheme; if you are in any doubt whatsoever (especially if several marks or candidates are involved) you should contact your Team Leader.

5. The following types of marks are available.

   M
   A suitable method has been selected and *applied* in a manner which shows that the method is essentially understood. Method marks are not usually lost for numerical errors, algebraic slips or errors in units. However, it is not usually sufficient for a candidate just to indicate an intention of using some method or just to quote a formula; the formula or idea must be applied to the specific problem in hand, eg by substituting the relevant quantities into the formula. In some cases the nature of the errors allowed for the award of an M mark may be specified.

A
Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. Accuracy marks cannot be given unless the associated Method mark is earned (or implied). Therefore M0 A1 cannot ever be awarded.

B
Mark for a correct result or statement independent of Method marks.

E
A given result is to be established or a result has to be explained. This usually requires more working or explanation than the establishment of an unknown result.

Unless otherwise indicated, marks once gained cannot subsequently be lost, eg wrong working following a correct form of answer is ignored. Sometimes this is reinforced in the mark scheme by the abbreviation isw. However, this would not apply to a case where a candidate passes through the correct answer as part of a wrong argument.

6. When a part of a question has two or more 'method' steps, the M marks are in principle independent unless the scheme specifically says otherwise; and similarly where there are several B marks allocated. (The notation 'dep *' is used to indicate that a particular mark is dependent on an earlier, asterisked, mark in the scheme.) Of course, in practice it may happen that when a candidate has once gone wrong in a part of a question, the work from there on is worthless so that no more marks can sensibly be given. On the other hand, when two or more steps are successfully run together by the candidate, the earlier marks are implied and full credit must be given.

7. The abbreviation ft implies that the A or B mark indicated is allowed for work correctly following on from previously incorrect results. Otherwise, A and B marks are given for correct work only — differences in notation are of course permitted. A (accuracy) marks are not given for answers obtained from incorrect working. When A or B marks are awarded for work at an intermediate stage of a solution, there may be various alternatives that are equally acceptable. In such cases, exactly what is acceptable will be detailed in the mark scheme rationale. If this is not the case please consult your Team Leader.

   Sometimes the answer to one part of a question is used in a later part of the same question. In this case, A marks will often be 'follow through'. In such cases you must ensure that you refer back to the answer of the previous part question even if this is not shown within the image zone. You may find it easier to mark follow through questions candidate-by-candidate rather than question-by-question.

8. Wrong or missing units in an answer should not lead to the loss of a mark unless the scheme specifically indicates otherwise.

   Candidates are expected to give numerical answers to an appropriate degree of accuracy. 3 significant figures may often be the norm for this, but this always needs to be considered in the context of the problem in hand. For example, in quoting probabilities from Normal tables, we generally expect *some* evidence of interpolation and so quotation to 4 decimal places will often be appropriate. But even this does not always apply – quotations of the standard critical points for significance tests such as 1.96, 1.645, 2.576 (maybe even 2.58 – but not 2.57) will commonly suffice, especially if the calculated value of a test statistic is nowhere near any of these values. Sensible discretion *must* be exercised in such cases.

Discretion must also be exercised in the case of small variations in the degree of accuracy to which an answer is given. For example, if 3 significant figures are expected (either because of an explicit instruction or because the general context of a problem demands it) but only 2 are given, loss of an accuracy ("A") mark is likely to be appropriate; but if 4 significant figures are given, this should not normally be penalised. Likewise, answers which are slightly deviant from what is expected in a very minor manner (for example a Normal probability given, after an attempt at interpolation, as 0.6418 whereas 0.6417 was expected) should not be penalised. However, answers which are *grossly* over- or under-specified should normally result in the loss of a mark. This includes cases such as, for example, insistence that the value of a test statistic is (say) 2.128888446667 merely because that is the value that happened to come off the candidate's calculator. Note that this applies to answers that are given as final stages of calculations; intermediate working should usually be carried out, and quoted, to a greater degree of accuracy to avoid the danger of premature approximation.

The situation regarding any particular cases where the accuracy of the answer may be a marking issue should be detailed in the mark scheme rationale. If in doubt, contact your Team Leader.

9. **Rules for crossed out and/or replaced work**

If work is crossed out and not replaced, examiners should mark the crossed out work if it is legible.

If a candidate attempts a question more than once, and indicates which attempt he/she wishes to be marked, then examiners should do as the candidate requests.

If two or more attempts are made at a question, and just one is not crossed out, examiners should ignore the crossed out work and mark the work that is not crossed out.

If there are two or more attempts at a question which have not been crossed out, examiners should mark what appears to be the last (complete) attempt and ignore the others.

NB Follow these maths-specific instructions rather than those in the assessor handbook.

10. Genuine misreading (of numbers or symbols, occasionally even of text) occurs. If this results in the object and/or difficulty of the question being considerably changed, it is likely that all the marks for that question, or section of the question, will be lost. However, misreads are often such that the object and/or difficulty remain substantially unaltered; these cases are considered below.

The simple rule is that *all* method ("M") marks [and of course all independent ("B") marks] remain accessible but at least some accuracy ("A") marks do not. It is difficult to legislate in an overall sense beyond this global statement because misreads, even when the object and/or difficulty remains unchanged, can vary greatly in their effects. For example, a misread of 1.02 as 10.2 (perhaps as a quoted value of a sample mean) may well be catastrophic; whereas a misread of 1.6748 as 1.6746 may have so slight an effect as to be almost unnoticeable in the candidate's work.

A misread should normally attract *some* penalty, though this would often be only 1 mark and should rarely if ever be more than 2. Commonly in sections of questions where there is a numerical answer either at the end of the section or to be obtained and commented on (eg the value of a test statistic), this answer will have an "A" mark that may actually be designated as "cao" [correct answer only]. This should be interpreted *strictly* – if the misread has led to failure to obtain this value, then this "A" mark must be withheld even if all method marks have been earned. It will also often be the case that such a mark is implicitly "cao" even if not explicitly designated as such.

On the other hand, we commonly allow "fresh starts" within a question or part of question. For example, a follow-through of the candidate's value of a test statistic is generally allowed (and often explicitly stated as such within the marking scheme), so that the candidate may exhibit knowledge of how to compare it with a critical value and draw conclusions. Such "fresh starts" are not affected by any earlier misreads.

A misread may be of a symbol rather than a number – for example, an algebraic symbol in a mathematical expression. Such misreads are more likely to bring about a considerable change in the object and/or difficulty of the question; but, if they do not, they should be treated as far as possible in the same way as numerical misreads, *mutatis mutandis*. This also applied to misreads of text, which are fairly rare but can cause major problems in fair marking.

The situation regarding any particular cases that arise while you are marking for which you feel you need detailed guidance should be discussed with your Team Leader.

Note that a miscopy of the candidate's own working is not a misread but an accuracy error.

11. Annotations should be used whenever appropriate during your marking.

The A, M and B annotations must be used on your standardisation scripts for responses that are not awarded either 0 or full marks. It is vital that you annotate standardisation scripts fully to show how the marks have been awarded.

For subsequent marking you must make it clear how you have arrived at the mark you have awarded.

12. For answers scoring no marks, you must either award NR (no response) or 0, as follows:

Award NR (no response) if:

- Nothing is written at all in the answer space
- There is a comment which does not in any way relate to the question being asked ("can't do", "don't know", etc.)
- There is any sort of mark that is not an attempt at the question (a dash, a question mark, etc.)

The hash key [#] on your keyboard will enter NR.

Award 0 if:

- There is an attempt that earns no credit. This could, for example, include the candidate copying all or some of the question, or any working that does not earn any marks, whether crossed out or not.

13. The following abbreviations may be used in this mark scheme.

M1	method mark (M2, etc, is also used)
A1	accuracy mark
B1	independent mark
E1	mark for explaining
U1	mark for correct units
G1	mark for a correct feature on a graph
M1 dep*	method mark dependent on a previous mark, indicated by *
cao	correct answer only
ft	follow through
isw	ignore subsequent working
oe	or equivalent
rot	rounded or truncated
sc	special case
soi	seen or implied
www	without wrong working

14. Annotating scripts. The following annotations are available:

✓and ✗	
BOD	Benefit of doubt
FT	Follow through
ISW	Ignore subsequent working (after correct answer obtained)
M0, M1	Method mark awarded 0, 1
A0, A1	Accuracy mark awarded 0, 1
B0, B1	Independent mark awarded 0,1
SC	Special case
^	Omission sign
MR	Misread

Highlighting is also available to highlight any particular points on a script.

15. The comments box will be used by the Principal Examiner to explain his or her marking of the practice scripts for your information. Please refer to these comments when checking your practice scripts.

    Please do not type in the comments box yourself. Any questions or comments you have for your Team Leader should be communicated by the *scoris* messaging system, e-mail or by telephone.

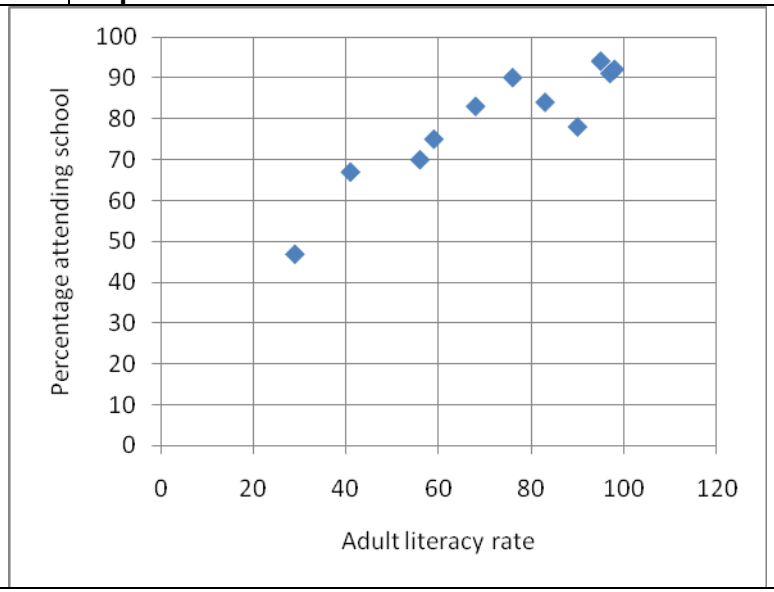
16. Write a brief report on the performance of the candidates. Your Team Leader will tell you when this is required. The Assistant Examiner's Report Form (AERF) can be found on the Cambridge Assessment Support Portal. This should contain notes on particular strengths displayed, as well as common errors or weaknesses. Constructive criticisms of the question paper/mark scheme are also appreciated.

17. Link Additional Objects with work relating to a question to those questions (a chain link appears by the relevant question number) – see scoris assessor Quick Reference Guide page 19-20 for instructions as to how to do this – this guide is on the Cambridge Assessment Support Portal and new users may like to download it with a shortcut on your desktop so you can open it easily! For AOs containing just formulae or rough working not attributed to a question, tick at the top to indicate seen but not linked. When you submit the script, *scoris* asks you to confirm that you have looked at all the additional objects. Please ensure that you have checked all Additional Objects thoroughly.

18. The schedule of dates for the marking of this paper is displayed under 'OCR Subject Specific Details' on the Cambridge Assessment Support Portal. It is vitally important that you meet these requirements. If you experience problems that mean you may not be able to meet the deadline then you must contact your Team Leader without delay.

Question	Expected Answer	Mark	Rationale/Additional Guidance	
1(i)	The pairing will eliminate any differences in individual subjects' aptitude in the word matching game and so will compare the before and after scores.	**E1** **E1**		Allow E1 for partially correct answer
(ii)	Differences are 14 -2 17 -5 -10 -7 15 4 9 19 Ranks of |d| are 7 1 9 3 6 4 8 2 5 10 Test statistic is 1+3+4+6 = 14 (or 2+5+7+8+9+10 = 41) Refer to paired Wilcoxon table with n=10 Lower 5% 1-tailed value is 10 Not significant Insufficient evidence to suggest that the median score after brain training is greater than the median score before brain training	**B1** **M1** **A1** **M1** **A1** **M1** **A1** **E1** **E1** **[9]**	No marks if differences not used FT if ranks wrong FT if previous M1 earned No FT if wrong No FT if wrong	No marks for two-sample test Do not award final E1 if mention of 'sufficient evidence to suggest no difference' or similar
(iii)	Population of differences in scores is Normally distributed. The t test take into account the actual values of the differences, rather than just the ranks, so is more powerful.	**E1** **E1** **E1** **E1** **[4]**	For population For differences For Normally distributed For more powerful	Allow 'underlying distribution' Allow 'more accurate' (BOD) Allow 'uses rankings rather than actual data so loses information'

Question	Expected Answer	Mark	Rationale/Additional Guidance	
2(i)	$\bar{x} = \frac{3308}{50} = 66.16$ $s^2 = \frac{222020 - \frac{3308^2}{50}}{49} = \frac{3162.72}{49} = 64.55$	**B1** **M1** **A1**	For S_{xx}	Allow even if over-specified as used later in the question
(ii)	Because both samples are large	**E1**		Condone 'sample is large'
(iii)	H_0: $\mu_W = \mu_M$ H_1: $\mu_W > \mu_M$ Where μ_W, μ_M denote the population mean times spent in the gym by women and men respectively 2–sample test based on N(0,1) Test statistic is $\frac{66.16 - 62.68}{\sqrt{\frac{64.55}{50} + \frac{134.0}{50}}} = \frac{3.48}{1.993} = 1.746$ 1–tailed 10% point of N(0,1) is 1.282 1.746 > 1.282 Significant There is evidence to suggest that on average women spend longer than men in the gym.	**B1** **B1** **B1** **E1** **M1** **M1** **A1** **B1** **M1** **A1** **E1** **[11]**	Condone absence of "population" if correct notation "μ" has been used, but do NOT accept $\bar{X}$ and $\bar{Y}$ or similar unless explicitly stated to be population means. Accept hypothesis explained in words, provided "population" appears. (Using correct version of test statistic) Numerator Good attempt at denominator CAO FT their test statistic provided both M1's earned No FT if CV wrong	

Question	Expected Answer	Mark	Rationale/Additional Guidance	
3(i)		**G1** **G1** **G1**	Linear axes, including labels Correct zero or clear broken scale All points correct (allow 2 errors)	
(ii)	The points do not appear to lie in an elliptical pattern so the population may not have a bivariate Normal distribution.	**E1** **E1**	For not elliptical For not bivariate Normal (allow opposite argument)	Allow E1 for mention of outliers
(iii)	<table><tr><td>Lit</td><td>41</td><td>56</td><td>90</td><td>97</td><td>83</td><td>95</td><td>98</td><td>29</td><td>59</td><td>76</td><td>68</td></tr><tr><td>Sch</td><td>67</td><td>70</td><td>78</td><td>91</td><td>84</td><td>94</td><td>92</td><td>47</td><td>75</td><td>90</td><td>83</td></tr><tr><td>R lit</td><td>2</td><td>3</td><td>8</td><td>10</td><td>7</td><td>9</td><td>11</td><td>1</td><td>4</td><td>6</td><td>5</td></tr><tr><td>R sch</td><td>2</td><td>3</td><td>5</td><td>9</td><td>7</td><td>11</td><td>10</td><td>1</td><td>4</td><td>8</td><td>6</td></tr><tr><td>d</td><td>0</td><td>0</td><td>-3</td><td>-1</td><td>0</td><td>2</td><td>-1</td><td>0</td><td>0</td><td>2</td><td>1</td></tr><tr><td>d^2</td><td>0</td><td>0</td><td>9</td><td>1</td><td>0</td><td>4</td><td>1</td><td>0</td><td>0</td><td>4</td><td>1</td></tr></table> $\Sigma d^2 = 20$ $r_s = 1 - \frac{6 \times 20}{11 \times 120} = 1 - 0.091 = 0.909$	**M1** **M1** **A1** **M1** **A1** **[5]**	For ranking (allow all ranks reversed) For d^2 For Σd^2 For method for r_s FT their ranks and Σd^2 provided $\lvert r_s \rvert < 1$ NB No ranking scores zero	

Question	Expected Answer	Mark	Rationale/Additional Guidance	
(iv)	H_0: no association between percentage attending school and adult literacy rate in the population H_1: positive association between percentage attending school and adult literacy rate in the population One tail test critical value at 1% level is 0.7091 Since 0.909 > 0.7091, there is sufficient evidence to reject H_0, i.e. conclude that there is enough evidence to suggest positive association between percentage attending school and adult literacy rate in the population.	**B1** **B1** **B1** **M1** **A1** **[5]**	For both hypotheses For population SOI NB H_0 H_1 <u>not</u> ito ρ For 0.7091 For comparison with c.v., provided $\lvert r_s \rvert < 1$ No FT if CV wrong For conclusion in words in context – FT their r_s and sensible cv	For two tail H_1 allow B1 for c.v. of 0.7545
4(i)	Treatment A is control Treatments B and C are experimental	**E1** **E1**	One mark if one error. No marks if 2 errors.	
(ii)	To check whether either fertiliser has an effect.	**E1**		
(iii)	Because there might be differences in the fertility of the three strips which would affect the results.	**E1** **E1**		
(iv)	Allocate numbers 1 to 100 to the plots. Use random numbers to choose 30 random numbers. If any repeats appear, choose further random numbers to replace them.	**E1** **E1** **E1**		Max E2 for 'numbering all plots from 1 to 100 and then selecting from a hat'

Question	Expected Answer	Mark	Rationale/Additional Guidance	
(v)	Assumptions required are: Normality of both populations, equal population variances.	 **B1** **B1**		 Do not allow 'samples are Normally distributed'
	H_0: $\mu_B = \mu_A$ H_1: $\mu_B > \mu_A$ Where μ_A, μ_B denote the population mean yields for treatments A and B	 **B1** **B1**	For both hypotheses	
	$\text{Pooled } s^2 = \frac{(7\times574.2)+(6\times612.7)}{13}$ = 591.97	**M1** **A1**	For attempt at pooling	Condone 8 and 7 in place of 7 and 6 for M mark
	Test statistic = $\frac{279-244.25}{\sqrt{591.97}\sqrt{\frac{1}{8}+\frac{1}{7}}} = \frac{34.75}{12.59} = 2.760$	**M1** **M1** **M1** **A1**	for numerator for $\sqrt{591.97}$ for $\sqrt{\frac{1}{8}+\frac{1}{7}}$ CAO	Condone 9 and 8 in place of 8 and 7 for M mark
	Refer to t_{13} 1–tail 5% point is 1.771	**M1** **A1**	Allow even if not all M marks earned	For two tail H_1 allow A1 for c.v. of 2.16
	Significant There is sufficient evidence to suggest that the population mean for peas grown with organic fertiliser is greater than that for peas grown with no fertiliser	**E1** **E1** **[14]**	FT from here if all M marks earned	
(vi)	H_0: the medians of the two populations are the same H_1: the medians of the two populations are different	**B1** **B1**	Allow 1 for medians Need population for second mark	
	Critical value = 38 42 > 38 so not significant Not enough evidence to suggest a difference between the population medians	**B1** **M1** **A1** **[5]**	 For not significant (dep on comparison)	No FT if CV wrong
Q4	**Total**	**[27]**		
	Total	**[72]**		

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

14 – 19 Qualifications (General)
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored

Oxford Cambridge and RSA Examinations
is a Company Limited by Guarantee
Registered in England
Registered Office; 1 Hills Road, Cambridge, CB1 2EU
Registered Company Number: 3484466
OCR is an exempt Charity

OCR (Oxford Cambridge and RSA Examinations)
Head office
Telephone: 01223 552552
Facsimile: 01223 552553

G243 Statistics 3

General Comments

The level of difficulty of the paper appeared to be appropriate for the candidates and there was no evidence of candidates being unable to complete the paper in the allocated time. Many candidates appeared to be well prepared for the paper although a significant number gained relatively few marks. In general, candidates supported their numerical answers with appropriate explanations and working, although in questions requiring discussion it was sometimes difficult to work out what the candidate meant.

Many candidates lost marks because they failed to mention 'population' in stating their hypotheses. Others lost marks because their answers were too assertive. Candidates should realise that they should express some doubt in the conclusion to a hypothesis using phrases such as 'there is sufficient (or insufficient) evidence to suggest that...'. Question 2 was found by most candidates to be rather easier than the other three questions, and in question 3(iii) it was disappointing to see the number of candidates who failed to rank the data. In question 4(iv) there were few good answers despite this being a straightforward question, simply requiring an explanation of how to select a random sample.

Comments on Individual Questions

1) Many candidates gained one mark in part (i), but few gave a fully correct response.

There were a good number of fully or almost fully correct responses to part (ii). However several candidates tried to carry out a two-sample test, gaining no marks. Other found the test statistic correctly but then got the critical value wrong. A very few candidates produced a fully correct solution other than their final conclusion.

In part (iii) most candidates mentioned 'Normal' in their answer but only a few mentioned 'population' and almost nobody mentioned 'differences'. All sorts of answers were seen for the reason for a *t* test being preferable, and many gained a generous mark, though none gave the actual response given in the mark scheme.

2) Part (i) was usually correct.

In part (ii) many candidates scored the mark which was allowed even for an answer such as *n* is large. However there was a variety of wrong answers, including 'samples random', 'Normally distributed', 'variances equal' etc.

In part (iii) there were few fully correct responses. Many candidates scored 2 marks out of 3 for the hypotheses, but missed the final mark for **'population** mean times'. It was pleasing to see that most candidates tried to carry out the correct test. The majority of candidates knew how to calculate the test statistic, although some squared the variances in the denominator. Most of these then went on to compare the test statistic with the correct critical value, but often missed the final mark because their conclusion was too assertive.

3) Part (i) was usually well answered, with most candidates gaining all 3 marks. However a significant number of candidates tried to plot a graph with 'country' on the horizontal axis and both of the other variables on the vertical axis.

In part (ii) many candidates mentioned 'an elliptical distribution' (or a non elliptical distribution) gaining one mark. Rather fewer mentioned 'bivariate Normality', also gaining one mark. Very few candidates mentioned both.

Part (iii) was reasonably well answered with many candidates ranking the data, although occasional slips in the ranking were seen. A number of candidates failed to rank the data and thus scored zero. Those candidates who got the ranks correct usually went on to get full marks for this part.

In part (iv) as in question 2, candidates often forgot to mention population and thus lost a mark. Some candidates gave a conclusion which was too assertive. Many candidates found the correct critical value but there was a wide variety of incorrect values.

4) Part (i) was generally well answered.

Many candidates gained the mark in part (ii), although the explanations differed in quality.

In part (iii) most candidates gained at least one mark, citing differences in fertility or difference in sunshine levels as possible confounding factors.

In part (iv), many candidates gave an explanation of systematic sampling, which gained no credit. Even those who tried to explain random sampling often only produced a partial explanation, often failing to mention repeats.

In part (v) many candidates realised that the populations need to be Normally distributed, but very few mentioned the need for equality of population variances. The hypotheses were usually stated correctly, although occasionally candidates gave two tailed hypotheses, but once again few mentioned population. Most candidates knew how to proceed with the hypothesis test, although a few made errors in the pooled variance or in calculating the test statistic. Most candidates who got this far were able to give the correct test statistic and complete the test, although a few tried to compare a negative test statistic to a positive critical value, thus losing the final two marks.

Candidates often omitted to mention 'median' in part (vi), thus losing a mark. Most then found the critical value correctly and usually went on to complete the test correctly, although a number thought that if the test statistic is greater than the critical value then the result is significant.

GCE Science (AS only)

			Max Mark	a	b	c	d	e	u
G641/01	Remote Sensing and the Natural Environment	Raw	60	45	40	35	31	27	0
		UMS	90	72	63	54	45	36	0
G642/01	Science and Human Activity	Raw	100	70	62	54	46	38	0
		UMS	150	120	105	90	75	60	0
G643/01	Practical Skills in Science	Raw	40	33	30	27	25	23	0
		UMS	60	48	42	36	30	24	0

GCE Sociology

			Max Mark	a	b	c	d	e	u
G671/01	Exploring Socialisation, Culture and Identity	Raw	100	73	66	59	53	47	0
		UMS	100	80	70	60	50	40	0
G672/01	Topics in Socialisation, Culture and Identity	Raw	100	71	63	55	47	39	0
		UMS	100	80	70	60	50	40	0
G673/01	Power and Control (A2)	Raw	100	70	62	54	47	40	0
		UMS	100	80	70	60	50	40	0
G674/01	Exploring Social Inequality and Difference (A2)	Raw	100	70	62	54	47	40	0
		UMS	100	80	70	60	50	40	0

GCE Spanish

			Max Mark	a	b	c	d	e	u
F721/01	Spanish: Speaking (AS): Externally Marked (OCR Repository)	Raw	60	46	41	36	31	26	0
F721/02	Spanish: Speaking (AS): Externally Marked (CD)	Raw	60	46	41	36	31	26	0
F721/03	Spanish: Speaking (AS): Visiting Examiner	Raw	60	46	41	36	31	26	0
F721	Spanish: Speaking	UMS	60	48	42	36	30	24	0
F722/01	Spanish: Listening, Reading and Writing 1	Raw	140	106	94	82	71	60	0
		UMS	140	112	98	84	70	56	0
F723/01	Spanish: Speaking (A2): Externally Marked (OCR Repository)	Raw	60	47	41	35	30	25	0
F723/02	Spanish: Speaking (A2): Externally Marked (CD)	Raw	60	47	41	35	30	25	0
F723/03	Spanish: Speaking (A2): Visiting Examiner	Raw	60	47	41	35	30	25	0
F723	Spanish: Speaking	UMS	60	48	42	36	30	24	0
F724/01	Spanish: Listening, Reading and Writing 2 (A2)	Raw	140	101	89	77	65	54	0
		UMS	140	112	98	84	70	56	0

GCE Statistics (MEI)

			Max Mark	a	b	c	d	e	u
G241/01	(Z1) Statistics 1	Raw	72	53	45	38	31	24	0
		UMS	100	80	70	60	50	40	0
G242/01	(Z2) Statistics 2	Raw	72	55	48	41	34	27	0
		UMS	100	80	70	60	50	40	0
G243/01	(Z3) Statistics 3	Raw	72	56	48	41	34	27	0
		UMS	100	80	70	60	50	40	0